measure of treatment satisfaction (6 items) and perceived frequency of hyper-/hypoglycaemia (2 items), validated in >100 languages. Arabic and French versions for Algeria were required for multinational clinical trials, necessitating linguistic validation of both languages. **METHODS:** FDA- and ISPOR-compliant methodology was used: 1) Review of developer-provided materials including concept definitions; 2) adaptation for Algeria from existing Arabic and French versions of the DTSQ (already validated for use elsewhere); 3) back translations into English and review by TransPerfect; 4) developer's team (HPR) review and agreement on interim version; 5) diabetes clinician review; 6) further HPR review; 7) cognitive interviewing with 5 patients with diabetes (Type 1 and 2)each for the two languages; 8) final HPR review; 8) final proofreading / formatting by native-language linguists. **RESULTS:** Adaptation of the Arabic and French language versions for Algeria presented interesting challenges. These included variable literacy levels, multiple dialects, as well as the complex relationships between vernacular Algerian Arabic and Modern Standard Arabic. Various changes to the standard French were required for comprehension to be ensured in Algeria. **CONCLUSIONS:** Modern Standard Arabic was judged to be the most appropriate form of Arabic for Algeria. While some minor changes were implemented to adjust for the local population, the wide variability throughout Algeria made full adaptation to a specific dialect inappropriate. Standard French without adaptation was found to be unsuitable for Algeria. Also notable is the importance of interviewing different respondents for the different languages (given many Algerians are fluent in both languages). Linguistically validated Arabic and French versions of the DTSQ for Algeria are now available for use.

PRM92

DESIGN OF A DISCRETE CHOICE EXPERIMENT FOR PREFERENCES ELICITATION OF ELDERLY POPULATION TOWARDS DEPRESSION

Monteiro AL[1], Santos MS[1], Cintra MACT[1], Pereira CC[2], Floriano DB[1], Plotkowski LM[3]
[1]*National Institute of Cardiology, Rio de Janeiro, Brazil,* [2]*Oswaldo Cruz Foundation, Rio de Janeiro, RJ, Brazil,* [3]*Centro de Estudos e Pesquisa do Envelhecimento, Rio de Janeiro, Brazil*

OBJECTIVES: Prioritize investment and tailoring interventions in the field of mental health care depends, largely, of the availability of decision-relevant data on population preferences. Preference elicitation methods, particularly, Discrete Choice Experiments (DCE) application has becoming increasingly popular in the field of health care research. The present study aims to describe the design of a DCE approach applied to elicit elderly preferences towards depression. **METHODS:** The set of attributes and levels was developed from a literature review and in consultation with experts. The PHQ-9 Depression scale was selected to base the wording of attributes and levels. The design of the DCE, and the software presentation were refined resorting to a pilot study, composed 20 blocks of choice tasks of 2 pairs of scenarios. Each scenario was composed by 5 attributes. To evaluate the respondent's performance during the DCE we relied on data on about time to complete the task, and difficulties in response as they were reported by the participants and observed by the interviewer. An electronic version of the instrument was developed by a team of programmers. **RESULTS:** A total of 10 elders, aged 60 to 82, completed the questionnaire. The overall understanding of the instrument was satisfactory, however all of them found fulfilling 20 choice tasks tiresome. The participants' performance improved with practice. **CONCLUSIONS:** The DCE was adjusted, according to a result of the pilot study. As major changes we sign the reduction of the number of attributes per option choice (from 5 to 4), the reduction of the number of choice tasks (from 20 to 10) and the introduction of warm-up choice tasks, as an example, to help the interviewee to understand the exercise logic. Application of DCE, using a Tablet PC as platform, was well tolerated by the elderly, and seems to be a viable alternative to other methods of extracting preferences alternative.

PRM93

COMPARING STANDARD GAMBLE AND VISUAL ANALOG SCALE MEASURES OF UTILITY FOR ERECTILE DYSFUNCTION HEALTH STATE

Bakken DG, Tomaszewski K, Wasserman D, Claeys C
KJT Group, Inc., Honeoye Falls, NY, USA

OBJECTIVES: Trade-off measures such as standard gamble (SG) generally are regarded as more valid than direct elicitation measures such as visual analog scale (VAS). The objective was to compare SG and VAS measures of utility for erectile dysfunction health states. **METHODS:** A representative (U.S.) sample of 1,000 adults completed an online survey about their own (for men) or partner's (for women in a heterosexual relationship) erectile function health states and quality of life. Health states utilities were measured using both SG and VAS questions (counterbalanced order). **RESULTS:** Looking at the within-subjects comparisons, SG and VAS measures yield very similar but not identical utility estimates for the erectile dysfunction ED health state. The mean utility estimates across all eligible respondents are 0.41 (0-1 scale) for the SG and 0.44 for the VAS. Looking just at SG, males have lower utility on average (0.37) for the ED health state than do females for their partner's ED health state. There was no effect due to the order of presentation. Additional analysis using covariate matching is underway to determine if there are between-subjects differences in utility for the two measures. **CONCLUSIONS:** With respect to obtaining estimates of health state utility or quality of life associated with ED, the SG and VAS measures yielded utility estimates that are comparable for decision-making purposes. The VAS is easier to administer than the SG, requiring as few as two questions. Pending the covariate matching to compare the utility estimates for the first position measure, the results indicate that VAS yields utility estimates for ED health state that are very close to estimates obtained using SG.

PRM94

INTEGRATING THE CHILD'S VOICE IN ADVERSE EVENT REPORTING IN ONCOLOGY TRIALS: INITIAL STEPS IN THE PEDIATRIC PRO-CTCAE INITIATIVE

Reeve BB[1], Withycombe J[2], Tomlinson D[3], Sung L[3], Mowbray C[4], Martens C[1], Mack JW[5], Lyons JC[1], Kohler RK[1], Keller S[4], Jacobs SS[4], Gold SH[1], Freyer DR[6], Baker JN[7], Hinds PS[4]
[1]*University of North Carolina at Chapel Hill, Chapel Hill, NC, USA,* [2]*Palmetto Health Children's Hospital, Columbia, SC, USA,* [3]*The Hospital for Sick Children, Toronto, ON, Canada,* [4]*Children's National Medical Center, Washington, DC, USA,* [5]*Harvard Medical School, Boston, MA, USA,* [6]*University of Southern California, Los Angeles, CA, USA,* [7]*St Jude Children's Research Hospital, Memphis, TN, USA*

OBJECTIVES: Over 60% of children with cancer will participate in a clinical trial. Collection of Adverse Events (AE) in trials is required by law and the current standard is that AEs are reported only by clinicians even when many AEs are subjective (e.g., fatigue, pain). Our overarching goal is to design a child self-report measure of subjective AEs that will inform AE reporting for the National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE), the standard lexicon for grading AEs in oncology trials world-wide. This study's objectives are to review the CTCAE to identify which AEs are amenable to self-report by children ages 7 or older and to design a self-report questionnaire to measure the selected AEs. **METHODS:** The CTCAE contains 790 AEs described in medical terms. Through two online surveys, we sought consensus among 184 experienced pediatric oncologists and nurses, from 7 diverse Children's Oncology Group treatment centers. To inform draft of the questionnaire, a search of PubMED was conducted for self-report symptom measures appropriate for children younger than 21 years. **RESULTS:** The response rates for the clinician surveys were 78% and 67%. A final 63 CTCAE terms were selected as being subjective, relevant for measuring in pediatric cancer trials, and amenable to self-report. The PubMED search yielded 28 self-report symptom measures for use in children with cancer. The measures varied in symptoms measured, characteristics (e.g., reference period, response options), and level of psychometric evidence. **CONCLUSIONS:** The 63 CTCAE terms were translated through this process into child friendly terms (e.g., "fatigue" to "tired") and a 129-item measure was drafted in English to capture a symptom's severity, frequency, and/or interference. Follow-up steps include cognitive interviews and a longitudinal validation study. Ultimately, systematic collection of these data in trials will enhance the validity of AE reporting and improve care for children with cancer.

PRM95

AN OPEN RESEARCH EXCHANGE FOR ONLINE PATIENT FEEDBACK IN PRO DEVELOPMENT

Harrington M, Heywood B, Rura S, Wicks P
PatientsLikeMe, Cambridge, MA, USA

OBJECTIVES: Patient-reported outcome (PRO) instrument development is a costly and lengthy iterative process. Access to patients, logistical difficulties, and the cost involved in getting patient feedback can delay the development of measures that adequately reflect patient experience. To address these challenges we built the Open Research Exchange (ORE) software platform. **METHODS:** ORE is integrated with PatientsLikeMe, an online community for patients who share their data for research and support. ORE integrates qualitative and quantitative stages of instrument feedback to support ongoing patient engagement throughout PRO development. Concept elicitation involves administration of open-ended questions while an equivalent to cognitive debriefing evaluates clarity, relevance, and adequacy of response options for each item. Larger-scale fielding alongside validated instruments permits psychometric validation. **RESULTS:** The effectiveness of the ORE platform was evaluated while developing the Insomnia Impact Questionnaire (IIQ). A survey of 11 open-ended questions was administered to 150 PatientsLikeMe members reporting mild to severe insomnia in the preceding 30 days. After 7 days of fielding, 76 complete surveys yielded 16,331 words of concept elicitation. Based on the obtained content, 31 questions were generated and administered during the cognitive interview phase to another sample of 150 PatientsLikeMe members matching above eligibility criteria. After 7 days, 59 surveys were completed. Feedback regarding the questions was obtained in the form of free text (6,934 words) and through a quantitative approach demonstrating that 90% patients reported high relevance, 95% reported easy understanding, and 95% reported adequate response option choice. **CONCLUSIONS:** The ORE online platform enables rapid and effective patient engagement in the PRO development process at a large scale.

PRM96

META-ANALYSIS OF PREFERENCE-BASED QUALITY OF LIFE VALUES IN HEART FAILURE

Stevanovic J[1], Pehlivanoglou P[2], Postma MJ[1]
[1]*University of Groningen, Groningen, The Netherlands,* [2]*Toronto Health Economics and Technology Assessment, Toronto, Canada*

OBJECTIVES: A large number of studies measuring quality of life (QoL) in heart failure (HF) patients is available in the literature. However, only the values measured with the preference-based instruments (e.g. EQ-5D UK and US "tariff" version, time trade-off (TTO), standard gamble (SG)) can be directly applied in a cost-utility analysis (CUA). Our aim was to summarize instrument-specific values in HF while accounting for study-level covariates and for the correlation between multiple values both within and between studies. **METHODS:** The instrument-specific values in HF were identified through a systematic search. Study-level covariates recorded from the eligible studies were: disease-severity level, publication year and mean population age. Finally, a multivariate meta-regression model was used to estimate pooled, instrument-specific values in HF. **RESULTS:** Our systematic search identified 22 studies reporting 28 values in HF. Instrument-specific estimates pooled on the base-case dataset were: 0.6385 (EQ-5D UK "tariff"), 0.7593 (EQ-5D US "tariff"), 0.6382 (SF-6D), 0.6534 (SG) and 0.7287 (TTO). Given that the EQ-5D (UK "tariff") was the most commonly applied instrument for measuring preference-based values among the included studies, the impact of study-level covariates was examined on those instrument-specific values. Age and disease-severity level were the study-level covariates that showed a significant impact on pooling QoL estimates. Yet, meta-regression modelling also identified a large amount of unexplained heterogeneity remaining. **CONCLUSIONS:** This is the first meta-analysis of QoL values in HF. Given the abundance of QoL measurements in HF and the requirement for applying single, accurate QoL values in a CUA, pooled estimates could be highly applicable in a CUA.